Technology-Enhanced Assessment and Feedback: How is evidence-based literature informing practice?

Denise Whitelock[1], Lester Gilbert[2], and Veronica Gale[2]

[1] Institute of Educational Technology, The Open University, UK

[2] Learning Societies Lab in the School of Electronics and Computing Sciences, University of Southampton

Abstract

This desktop research commissioned by the Higher Education Academy set out to consult with the academic community about which references on assessment and feedback with technology enhancement were most useful to practitioners. While all the recommended publications may be characterised as reputable and the majority were peer-reviewed (67.7%), only a minority provided quantitative data (28.2%), of which relatively few provided appropriate experimental designs or statistical analysis (18.5%). The majority of publications were practitioner-led case studies. The references that were recommended to us are clearly having an impact on current practice and are found valuable by practitioners. The key messages from these sources are consistent and often give detailed and practical guidance for other academics. We found that most of the recommended literature focused on the goals that technology enhancement can enable assessment and feedback to meet and how assessment and feedback can be designed to make best use of the technology.

Introduction

One of the main drivers for Learning has long been acknowledged as the Assessment that students must undergo during the course of their studies (Rowntree, 1987). The "backwash effect" of assessment (Biggs, 1996), such as "students only learn what is assessed", highlights how the results influence certain assessment practices. There is now a sea change in attitudes to assessment where quite rightly the role of assessment is focused to support student learning (Assessment for Learning Group, 2002). Providing students with constructive, timely and "easy to understand" feedback is taking centre stage in this new culture of assessment (Havnes & McDowell, 2008) and has gained increased interest throughout the HE Sector with the advent of electronic assessment.

There is now a growing body of literature about the role that electronic assessment and feedback is playing in the HE Sector and one of the salient questions of the day is how is this literature informing practice to support student learning?

This paper reports on a study commissioned by the Higher Education Academy in January 2010 undertaken to investigate this question through addressing the following:

1. Consult the academic community about which references on assessment and feedback with technology enhancement are most useful to practitioners

2. Prioritise evidence-based references i.e. those that are peer reviewed and have data to support their practice

3. Synthesise the main points from these references

4. Provide signposts for readers to locate the original documents for further study

The aim of this desktop research was to support the Higher Education Sector in its use of technology to enhance learning and teaching by providing a comprehensive and useful synthesis of evidence based practice in this domain.

Methodology

The literature on assessment and feedback with technology enhancement is large and varied. In order to focus on references that practitioners find particularly helpful and would recommend to their peers, we consulted the HE community via:

- An Advisory Panel

- An Assessment seminar series organized by the Academy

- A stand at the CAA 2010 Conference

- Email requests to an 'expert group' of HE practitioners

We employed a paper questionnaire and an online survey powered by Survey Monkey to elicit the information from participants at the Academy's seminar series and at the CAA 2010 Conference. The community was also alerted to this study via Cloudworks (http://cloudworks.ac.uk/cloud/view/2952).

We received 142 references, including journal articles, reports, books and websites. These were reviewed to identify:

- The technology-enhanced methods discussed (as opposed to assessment and feedback in general)

- The use of technology for assessment and feedback (rather than other aspects of student learning, course administration, or content management)

- The type of evidence that was provided to support their findings or observations

- How easy the reference was to access (we were unable to source 18 references, particularly those published within a specialist group, a long time ago, or in another country)

The examples included in this desktop research adhered to the following definition for 'Technology enhancement'. In this study this was any method that involved a computer or other technology (such as podcasting) in providing advice before the assessment, during the setting of the assessment activity, supporting the assessment activity, capturing student responses, and providing generic or individual feedback on performance. Examples of the technologies discussed in the recommended texts include audio and video feed forward and feedback; e-portfolios; tests delivered, answered and marked by computers; electronic voting systems; web-enabled group discussions; and peer reviews.

Technology use that was excluded from the project included online submission of written work and computer collection of marks.

'Technology enhancement' suggests that using the technology provides better quality than the alternative, perhaps paper-based materials or a lecture. Although a case can be made that technology does enhance learning and teaching quality in some cases, technology 'enhancement' may also be considered in terms of cost savings or productivity improvements.

Findings from the evidence-based literature

The definition for evidence-based literature that was used for this study was research/investigations supported by data, including validity and reliability measures, comparisons of learning achieved with and without technology-enhanced features, effect size estimates, and quantified estimates of time saved or effort needed. Review, 'compilation', or meta-study references were considered 'evidence-informed' and were treated in the same way as evidence-based references.

We developed and refined five categories of grades of evidence (see Table 1). In general, a study meeting a category listed earlier in the table can also match categories listed later.

References to studies were allocated to a particular category based on the information provided in the article or report. When a case study was cited in a 'compilation', meta-study, or review report we tried to find the underlying source but were not always successful. The categories do not imply a hierarchy of value to the reader.

Technology Enhanced Assessment and Feedback: How is evidence-based literature informing practice?

Table 1. Categories of evidence used in this report

Category	Description
1a	Peer-reviewed generalizable study providing effect size estimates and which includes (i) some form of control group or treatment (may involve participants acting as their own control, such as before and after), and / or (ii) blind or preferably double-blind protocol.
1b	Peer-reviewed generalizable study providing effect size estimates, or sufficient information to allow estimates of effect size.
2	Peer-reviewed 'generalizable' study providing quantified evidence (counts, percentages, etc) short of allowing estimates of effect sizes.
3	Peer-reviewed study.
4	Other reputable study providing guidance.

The categories of evidence outlined above may be applied equally to different kinds of study which target a variety of research questions.

Table 2. Number of references recommended in each evidence category

| Evidence category	Number of references recommended[a]	Cumulative %
1a	15	12.1%
1b	8	18.5%
2	12	28.2%
3	49	67.7%
4	40	100.00%
Total	124	

It is interesting to note that the majority of papers recommended by the practitioners belong to category 3 which consisted of peer reviewed studies. However the recommended references which belonged to category 1a were disappointingly fewer than expected. 25% of these referred to work undertaken with school children. A number of other papers clustered around the use of e-portfolios (see Barbera, 2009; Chang & Tseng, 2009; Loddington et al, 2009). Other papers in this category focused around e-assessment in specific subject domains such as health education (Lee & Weerakoon, 2001), medical knowledge (Mitchell et al, 2003), mathematics (Sandene et al, 2005). More importantly if electronic feedback is to influence learning then it should affect self regulation and two papers in this category reported work on this facet of assessment, see Strang (2010) and Nicol (2009). Strang's findings revealed that students from collectivist and risk-taking cultures were more likely to obtain higher grades because they were willing to use deep or strategic study approaches when evaluating the e-feedback. Deep and strategic study approaches were more effective in processing e-feedback.

Designing assessment and feedback with technology-enhancement

Assessment and feedback forms an integral component of any learning design and is modeled in Figure 1 below. This figure (Gilbert & Gale, 2007) illustrates that an 'atomic' learning transaction fosters a learning/teaching dialogue which should be linked to learning objectives.

Figure 1. The Instructional Transaction
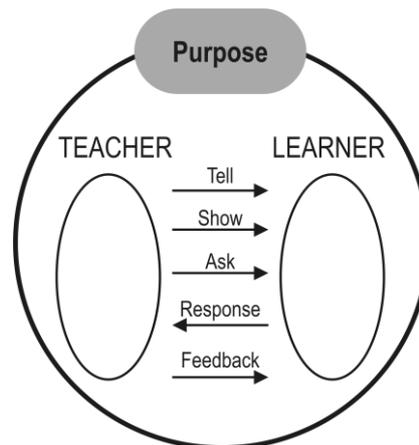
Some examples of learning designs that use technology enhancement successfully

An example of how a successful learning design can be enhanced by technology is provided in Crouch and Mazur (2001). Their paper describes the results of ten years' experience of improved student results (compared with traditional instruction and therefore in evidence category 1b) using a method they call Peer Instruction:

> "A class taught with PI [Peer Instruction] is divided into a series of short presentations, each focused on a central point and followed by a related conceptual question [MCQ example given]. Students are given one or two minutes to formulate individual answers and report their answers [using a poll] to the instructor. Students then discuss their answers with others sitting around them; the instructor urges students to try and convince each other of the correctness of their own answer by explaining the underlying reasoning. Finally, the instructor […] polls students for their again (which may have changed based on the discussion), explains the answer and moves onto the next topic."

They found that the "vast majority" of students who changed their vote after the peer discussion moved from an incorrect answer to the correct answer.

Draper (2009) discusses how this technique can be used with an electronic voting system, a technology used to display the question, capture the student responses, and display the votes for each option as a graph.

Technology enhancement is not just applied to MCQs, however. Jordan and Mitchell (2009) (category 1a) provide evidence for moving beyond the MCQ and using open questions with technology enhancement. They suggest that open questions are

suitable for computerised delivery and feedback "if correct answers can be given in short phrases or simple sentences and the difference between correct and incorrect answers is clear-cut." Whitelock & Watt (2008) illustrate this using the Open University's 'Open Comment' system.

The lower levels of Bloom's taxonomy are where many academics start to apply technology-enhanced assessment methods. However, there is evidence in the literature of how academics have moved into higher level learning outcomes and more subtle assessment questions when applying technology enhancement.

Ashton et al (2006) provide evidence (category 1a) that technology-enhanced methods can be used to mirror tutor marking practices in mathematical examinations. They explain how software was developed and some questions were redesigned to allow partial credits to be awarded and mathematical expressions to be entered by students in automated exams. This work was undertaken as part of the PASS-IT project.

Boyle and Hutchinson (2009) address the issue of whether or not sophisticated tasks can be assessed using technology enhancement. In this paper (category 3) they suggest that, "e-assessment will become an important and widely-used feature of education systems in the near future. Further, the types of questions and tasks used in near-future e-assessment may well be quite different from questions and tasks used in on-paper assessment, and in early implementations of computerised assessment."

The case for developing appropriate feedback

When Maclellan (2001) conducted a survey (evidence category 2) of staff and students at the University of Strathclyde on their perceptions of assessment for learning, students' perceptions were that the feedback they received was not "routinely helpful in itself or a catalyst for discussion."

This finding is supported by the results of the National Union of Students (NUS) surveys and the NUS emphasise the fact that "effective feedback on assessment is a crucial aspect of assessment processes and a key feature of enhancing the learning process."

Further evidence for the role of feedback in an assessment strategy is provided by Whitelock (2010, evidence category 4) who points out that, "Formative assessment assists the on-going learning cycle while summative assessment is not cyclical and assists with ascertaining the progress made by students at a particular moment in time e.g., at the end of a course." Feedback from frequent formative assessment is therefore a vital component of a course's learning strategy.

Specific sets of feedback guidelines have been produced by Gibbs and Simpson (2004) and Nicol & Macfarlane-Dick (2006). While further evidence that information to move students from their current to the desired performance is the key for the effective use of e-feedback is provided by Dexter (2010). In this category 3 paper Dexter presents some findings from the design and evaluation of software known as ETIPS (educational theory into practice software). It was developed to provide K-12 educators with case studies where they could practise using concepts from their university education courses in a classroom and school setting. A "student model" is used which defines the declarative, procedural and contextual knowledge that students are being tested on with the e-feedback explaining how the student can improve their performance. It is vital that students perceive that the knowledge and

skill captured in the student model is important, so that they respond to the e-feedback with further learning.

Designing audio feedback

As well as the general principles of feedback provided above, the references recommended to us included specific advice for designing audio feedback. Middleton and Nortcliffe's (2010) study was designed to identify factors which affect the implementation of audio feedback. Semi-structured interviews were conducted with academics about their use of audio feedback (evidence category 3) and the authors have produced a set of audio feedback design principles from their findings. These principles are that audio feedback should be:

1. *"Timely and meaningful.*

2. *Manageable for tutors to produce and the learner to use.*

3. *Clear in purpose, adequately introduced and pedagogically embedded.*

4. *Technically reliable and not adversely determined by technical constraints or difficulties.*

5. *Targeted at specific students, groups or cohorts, addressing their needs with relevant points in a structured way.*

6. *Produced within the context of local assessment strategies and in combination, if appropriate, with other feedback methods using each medium to good effect.*

7. *Brief, engaging and clearly presented, with emphasis on key points that demand a specified response from the learner.*

8. *Of adequate technical quality to avoid technical interference in the listener's experience.*

9. *Encouraging, promoting self esteem.*

10. *Formative, challenging and motivational."*

Designing feedback: when to give it

The best time to give feedback is when there is still time for the student to act upon it and improve their performance (Nicol and MacFarlane-Dick, 2006 and Sadler 1989). With some technology-enhanced methods immediate feedback is intrinsic to the skill being assessed, for example through computer simulation of the task.

Nicol and MacFarlane-Dick (2006) suggest the following strategies:

- *"Provide feedback on the work in progress and increase opportunities for resubmission.*
- *Give feed forward and group feedback.*
- *Set two-stage assignments: provide action points then involve students in groups identifying their own action points."*

Although tutors might offer feedback at a time when students can act upon it to improve their performance, students may not take the opportunity to do so unless encouraged. Whitelock (2010) cites previous authors (Wojtas, 1998; Weaver, 2006; Price and O'Donovan, 2008) who found that, although students ask for feedback

they often seem only interested in their grade. Lindsay Jordan reports on efforts to engage students with all aspects of their feedback:

"Asking students to respond to their feedback has been found to be a highly effective method of increasing distance learners' engagement with summative feedback. Feedback response tasks encourage students to focus on how they will use their feedback in the future, while simultaneously allowing them to enter into a dialogue with the tutor on any points that require clarification." Jordan (2009, evidence category 3.)

Using peer assessment to give feedback

A feedback approach that is likely to engage students is to include peer assessment. Nicol (2010, evidence category 3) discusses how peer assessment can be used to provide "feedback dialogue" and puts forward the following benefits:

"Peer feedback scenarios where students receive comments on an assignment from many other students provide a richness and volume of dialogue that is difficult for a single teacher to match. In such situations, students must actively process and reprocess feedback input from a variety of sources and are potentially exposed to multiple levels of analysis and scaffolding.

[The] construction of feedback is likely to heighten significantly the level of student engagement, analysis and reflection with feedback processes.

[Further,] where peers generate and receive feedback in relation to the same assignment task (i.e. an essay that all students are writing), they learn not only about their own work but also about how it compares with productions of other students."

Nicol recognises that some students have a "lack of confidence in their peers and prior predispositions to solo working" and suggests that teachers comment on the peer comments when the peer working is first introduced to overcome these obstacles.

Key messages

To highlight the main findings from this desktop research, technology can be used to design assessment and feedback that:

- o Is more authentic to the skills being assessed;
- o Can assess a range of skills and provide accurate and helpful feedback in a short period of time using designs such as hints, partial credits and asking the student how confident they are of their answer;
- o Is more accurate than alternative methods, for example using electronic voting instead of a show of hands;
- o Adds 'pedagogical power' to multiple-choice question, for example by using assertion-reason questions;
- o Asks open questions;
- o Meets the learning needs of the contemporary learner who prefers: active, personalised and just in time learning; authentic tasks;

knowing where to search for information rather than memorising it; skilled use of tools and collaboration amongst other characteristics. This 'Assessment 2.0' makes use of the Internet, especially Web 2.0 to match assessment and feedback to the learning characteristics of students;

o Encourages peer and self-assessment, student learning, and reflection through a process of e-portfolios. This design can also be used to reduce the workload placed on the tutor for marking and giving feedback.

Discussion

The references that were recommended to us are clearly having an impact on current practice and are found valuable by practitioners. The key messages from these sources are consistent and often give detailed and practical guidance for other academics. We found that most of the recommended literature focused on the goals that technology-enhancement can enable assessment and feedback to meet and how assessment and feedback can be designed to make best use of the technology.

While all the recommended publications may be characterised as reputable (Table 2) and the majority were peer reviewed (67.7%), only a minority provided quantitative data (28.2%), of which relatively few provided appropriate experimental designs or statistical analysis (18.5%). The majority of publications proposed to us were practitioner-led case studies. Few of the publications were based on evidence categories 1a, 1b and 2 (Table 1).

Most recommended publications lacked direct application. This is because the majority of publications proposed to us were case studies which were highly contextualised to the specifics of a course or organisation and any generalisations were usually at an exceptionally high level.

On the basis of this research we suggest that:

• Risk and change management strategies are introduced when committing significant resources.

This is because the majority of references are largely case studies and lack quantified outcomes, prudence may be needed when changing to technology-enhanced assessment strategies at course or institutional level.

• Learning is addressed first while the affordances of technology enhancement take second place.

In engaging with technology-enhanced learning, the majority of references concluded that success followed from addressing the learning needs and not shoehorning the learning into the technological tools.

Acknowledgements

References

ARG (2002). *Assessment for Learning: 10 principles*. Retrieved May 20, 2011 from assessment-reform-group.org

Ashton, H. S., Beevers, C. E., Korabinski, A. A. & Youngson, M. A., (2006). Incorporating partial credit in computer-aided assessment of Mathematics in secondary education. *British Journal of Educational Technology*, 37(1), 93–119.

Barbera, E. (2009). Mutual feedback in e-portfolio assessment: an approach to the netfolio system. *British Journal of Educational Technology,* 40(2), 342-357.

Biggs, J. (1996). Assessing learning quality: Reconciling institutional, staff and educational demands. Assessment and Evaluation in Higher Education, 21(1), 5-15.

Boyle, A. & Hutchison, D. (2009). Sophisticated tasks in e-assessment: what are they and what are their benefits? *Assessment & Evaluation in Higher Education,* 34(3), 305-319.

Chang, C. & Tseng, H. (2009). Use and performances of web-based portfolio assessment. *British Journal of Educational Technology,* 40(2), 358-370.

Crouch, C. H. & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics,* 69, 970-977.

Dexter, S. (2010). E-feedback intersections and disconnections in the interests of designers and users. In D. Whitelock & P. Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning,* 20(2), 169-188.

Draper, S. (2009). Catalytic assessment: understanding how MCQs and EVS can foster deep learning. *British Journal of Educational Technology,* 40(2), 285-293.

Gibbs, G. & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education,* 1, 3–31.

Gilbert, L., Whitelock, D. & Gale, V. (In Press). *Synthesis Report on Assessment and Feedback with Technology Enhancement.* Technical Report, University of Southampton.

Gilbert, L. & Gale, V. (2007). *Principles of E-Learning Systems Engineering*. Chandos Publishing (Oxford) Ltd.

Havnes, A. & McDowell, L. (2008) *Balancing Dilemmas in Assessment and Learning in Contemporary Education.* New York: Routledge.

Jordan, L (2009). Transforming the student experience at a distance: designing for collaborative online learning. *Engineering Education*, 4(2), 25-36.

Jordan, S. & Mitchell, T. (2009). e-assessment for learning? The potential of short-answer free-text questions with tailored feedback. *British Journal of Educational Technology,* 40(2), 371-385.

Lee, G. & Weerakoon, P. (2001). The role of computer-aided assessment in health professional education: a comparison of student performance in computer-based and paper-and-pen multiple-choice tests. *Medical Teacher,* 23(2), 152-157.

Loddington, S., Pond, K., Wilkinson, N. & Willmot, P. (2009). A case study of the development of WebPA: An online peer-moderated marking tool. *British Journal of Educational Technology,* 40(2), 329-341.

MacLellan, E. (2001). Assessment for Learning: The differing perceptions of tutors and students, *Assessment & Evaluation in Higher Education,* 26(4), 307-318.

Middleton, A. & Nortcliffe, A. (2010). Audio feedback design: principles and emerging practice. In D. Whitelock & P. Brna (eds) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning,* 20(2), 208-223.

Mitchell, T., Aldridge, N., Williamosn, W. & Broomhead, P. (2003*). Computer based testing of medical knowledge.* 7th International CAA Conference, Loughborough. Last accessed September 2010 from http://www.caaconference.com/pastconferences/2003/index.asp

Nicol, D. (2009). Assessment for learner self-regulation: enhancing achievement in the first year using learning technologies. *Assessment and Evaluation in Higher Education*, 34(3), 335-352.

Nicol, D. (2010). From monologue to dialogue: improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education,* 35(5), 501-517.

Nicol, D. J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education,* 31(2), 199-218.

Rowntree, D. (1987). *Assessing students: how shall we know them?* London: Kogan Page.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science,* 18(2), 119-144.

Sandene, B., Horkay, N., Bennett, R.E., Allen, N., Braswell, J., Kaplan, B. & Oranje, A. (2005). *Online assessment in Mathematics and Writing: Reports from the NAEP Technology-Based Assessment Project, Research and Development Series*. Last accessed August 2010 from http://nces.ed.gov/nationalreportcard/pubs/studies/2005457.asp

Strang, K.D. (2010). Measuring self regulated e-feedback, study approach and academic outcome of multicultural university students. In D. Whitelock & P. Brna (eds.) Special Issue 'Focusing on electronic feedback: feasible progress or just unfulfilled promises?' *Int. J. Continuing Engineering Education and Life-Long Learning*, 20(2), 239-255.

Whitelock, D. (2010). Activating Assessment for Learning: are we on the way with Web 2.0? In M.J.W. Lee & C. McLoughlin (Eds.) *Web 2.0-Based-E-Learning: Applying Social Informatics for Tertiary Teaching.* IGI Global.

Whitelock, D. & Watt, S. (2008). Reframing e-assessment: adopting new media and adapting old frameworks. *Learning, Media and Technology,* 33(3) 153–156 Routledge, Taylor & Francis Group.